

DIAGONALS IN 2D-TILINGS AND COINCIDENCE DENSITIES OF SUBSTITUTIONS

F. MICHEL DEKKING

ABSTRACT. We study the diagonals of two-dimensional tilings generated by direct product substitutions. The properties of these diagonals are primarily determined by the eigenvalues of the substitution matrix, but also the order of the letters in the substitution plays a role. We show that the diagonals may fail to be uniformly recurrent, and that the frequencies of letters on the diagonal may not exist. We also highlight the connection with the density of coincidences and overlap distributions.

Key words. Direct product substitution; tiling; coincidence density; overlap distribution; self-similarity; substitution of Salem type.

MSC: Primary: 52C23; Secondary: 52C20, 37A25

1. INTRODUCTION

The two-dimensional tilings generated by direct product substitutions have a very simple structure, but we will show that the *diagonals* of these tilings are not that simple. Throughout the paper θ denotes a substitution (morphism) on $\{0, 1\}$ that admits two fixed points: we will require that the first letter of $\theta(0)$ is 0 and the first letter of $\theta(1)$ is 1.

The (*direct*) *product substitution* $\theta \times \theta$ of θ with itself is defined by

$$\theta \times \theta(i, j) = \theta(i) \times \theta(j), \quad i, j \in \{0, 1\},$$

where for two words v and w of length m and n we define the product word as the matrix

$$v \times w = ((v_k, w_\ell))_{1 \leq k \leq m, 1 \leq \ell \leq n}.$$

For example, when θ is the Thue-Morse substitution given by $\theta(0) = 01$, and $\theta(1) = 10$ then,

$$\theta \times \theta((0, 0)) = \begin{pmatrix} (0, 1) & (1, 1) \\ (0, 0) & (1, 0) \end{pmatrix}, \quad \text{and} \quad \theta \times \theta((0, 1)) = \begin{pmatrix} (0, 0) & (1, 0) \\ (0, 1) & (1, 1) \end{pmatrix}.$$

It is visually more convenient to represent this with 1×1 colored tiles:

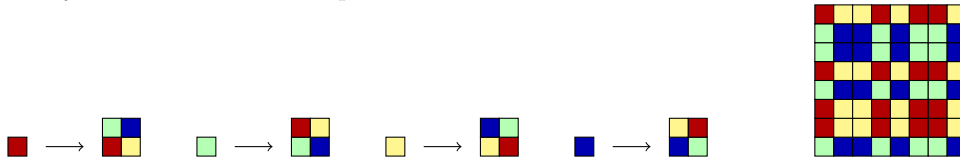


FIGURE 1. The Thue Morse direct product substitution, and an order 3 iterate.

Let $u^* = \theta^\infty(0)$ and $v^* = \theta^\infty(1)$ be the two fixed points of θ . Obviously the 2D-substitution $\theta \times \theta$ has four fixed points, $u^* \times u^*$, $v^* \times v^*$, $u^* \times v^*$, and $v^* \times u^*$. The first two 2D infinite words have trivial diagonals: these are just a doubling of u^* , respectively v^* . However, even though the diagonal of $u^* \times v^*$ is also generated by a substitution of length q when θ is a substitution of constant length q , this diagonal is not always so regular as in the $u^* \times u^*$ case.

From now on, we will often denote the letters of the alphabet $\{0, 1\} \times \{0, 1\}$ as column vectors, for easier reading. It is easy to see that the diagonal $\Delta^* = \Delta^*(0, 1)$ of the $u^* \times v^*$ fixed point of the Thue-Morse direct product substitution is generated by the substitution

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix}; \quad \text{in tiles:} \quad \begin{matrix} \text{green} & \rightarrow & \begin{matrix} \text{green} & \text{yellow} \\ \text{green} & \end{matrix} \\ \text{yellow} & \rightarrow & \begin{matrix} \text{yellow} & \text{green} \\ \text{yellow} & \end{matrix} \end{matrix}.$$

We see that the successive first coordinates of the diagonal will form u^* and the successive second coordinates of the diagonal will form v^* . This will always be the case as soon as θ is *continuous*, which means that $\theta(1) = \overline{\theta(0)}$, where we use the mirror substitution given by $\overline{0} = 1$ and $\overline{1} = 0$. When θ is not continuous, it is called *discrete*. The terminology follows [2]. As an example of a discrete substitution take

$$\theta(0) = 011 \quad \theta(1) = 101.$$

The substitution that generates the diagonal is then given by

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} 1 \\ 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

It follows that for discrete substitutions the diagonal is less regular: the sequence Δ^* is still recurrent, but no longer uniformly recurrent. This is caused by the occurrence of longer and longer words consisting solely of the letters $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. This also leads to the property that $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ have frequency 0 and $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ have a frequency equal to the original frequencies of 0 and 1 in the fixed point u^* of θ (cf. Lemma 4). The technical explanation of this phenomenon is that for discrete substitutions the substitution that generates the diagonal is no longer primitive, but reducible.

In Section 2 we consider the diagonal problem for substitutions of *non-constant* length. It turns out that in some cases a description of the diagonal is possible with the aid of the notion of *balanced blocks*, which has been used in the literature to compute coincidence densities. In Section 3 we extend the notion of balanced blocks, so that it is applicable to a larger class of non-constant substitutions. In Section 4 we show how to compute coincidence densities, and give an example where the coincidence density (and so the frequencies of letters) does not exist. In Section 5 we present a more positive result: using the same substitution to generate a self-similar tiling, the frequencies of letters on the diagonal do exist. In fact this is shown to be generally true for substitutions with rational eigenvalues. In Section 6 we discuss the case of two eigenvalues with modulus larger than 1. In Section 7 we discuss the connection with the notion of overlap distribution from statistical physics.

2. BALANCED BLOCKS AND THE DENSITY OF COINCIDENCES

Non-constant length substitutions can also generate tilings ([12]). The usual example is the Fibonacci substitution $0 \mapsto 01, 1 \mapsto 0$. But since this substitution does not have two fixed points, we will in the following rather consider its nephew given by

$$\theta : \quad 0 \mapsto 001, \quad 1 \mapsto 10.$$

If we want to generate the diagonal Δ^* of the corresponding $u^* \times v^*$ we now have the problem that the occurrences of $\theta(0)$ and $\theta(1)$ in u^* and v^* are no longer synchronized. The solution to this problem stems from [11], and is known as the *Method of balanced blocks*.

Let $N_i(A)$ denote the number of times letter i occurs in the word A . A ‘double’ word $\begin{pmatrix} A \\ B \end{pmatrix}$ is called a *balanced* block if

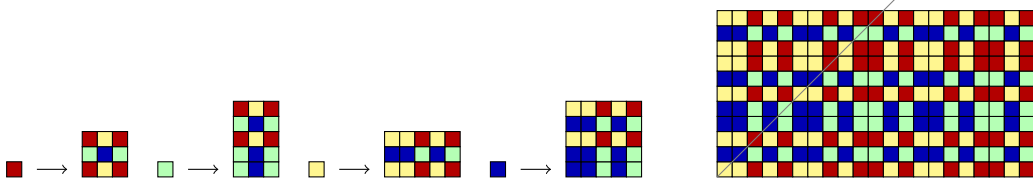
$$N_0(A) = N_0(B) \text{ and } N_1(A) = N_1(B).$$

The idea is to cut up the ‘double’ sequence $\Delta^* = \begin{pmatrix} u^* \\ v^* \end{pmatrix}$ into irreducible balanced blocks:

$$\begin{array}{l} u^* = |001|0|01|10|0|0|1|001|10|\dots \\ v^* = |100|0|10|01|0|0|1|100|01|\dots \end{array}$$

The key observation is that for any balanced block $\begin{pmatrix} A \\ B \end{pmatrix}$ the block $\begin{pmatrix} \theta(A) \\ \theta(B) \end{pmatrix}$ is obviously again balanced. This makes it possible to define a substitution $\hat{\theta}$ on blocks $\hat{A} = \begin{pmatrix} A \\ B \end{pmatrix}$ by

$$\hat{\theta}(\hat{A}) = \hat{A}_1 \dots \hat{A}_m,$$

FIGURE 2. Direct product of $\theta : 0 \mapsto 010, 1 \mapsto 11010$, and second iterate of $\binom{1}{0}$.

where $\hat{A}_1 \dots \hat{A}_m$ is the decomposition of \hat{A} into irreducible balanced blocks. For example, taking the first balanced block in the decomposition above:

$$\hat{\theta} : \begin{bmatrix} 001 \\ 100 \end{bmatrix} \mapsto \begin{bmatrix} 001 \\ 100 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} \begin{bmatrix} 01 \\ 10 \end{bmatrix} \begin{bmatrix} 10 \\ 01 \end{bmatrix}.$$

Continuing in this way, and identifying the blocks $\begin{pmatrix} A \\ B \end{pmatrix}$ and $\begin{pmatrix} B \\ A \end{pmatrix}$, one obtains $\hat{\theta}$ defined on the alphabet of (identified) irreducible balanced blocks

$$\hat{I} = \left\{ \begin{bmatrix} 001 \\ 100 \end{bmatrix}, \begin{bmatrix} 01 \\ 10 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}.$$

As in [11] (see also Lemma 4) it then follows that Δ^* behaves as in the case of discrete substitutions of constant length: the symbols $\hat{0} := \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and $\hat{1} := \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ dominate.

Actually [11] focusses on the so called *density of coincidences*, defined by

$$\delta_c := \lim_{n \rightarrow \infty} \frac{1}{n} \text{Card}\{1 \leq k \leq n : u_k^* = v_k^*\}.$$

Since counting coincidences is equivalent to counting occurrences of $\hat{0}$ and $\hat{1}$, the result is that $\delta_c = 1$ for this substitution. This plays an important role in the determination of the spectra of dynamical systems generated by the substitutions, see e.g., [11], and [7]. Crucial there is that θ is a substitution of Pisot type, i.e., the second eigenvalue of M_θ lies in the unit circle. In [1] it is shown that this implies that the number of irreducible balanced blocks is finite.

We return to our original question: what is the structure of the diagonal of a direct product substitution? As an example of a non-Pisot we consider a substitution of Salem type, i.e., the second eigenvalue lies on the unit circle. We take

$$\theta : 0 \mapsto 010, 1 \mapsto 11010.$$

Here the ‘double’ fixed point sequence $\Delta^* = \begin{pmatrix} u^* \\ v^* \end{pmatrix}$ starts with the irreducible balanced block $\begin{bmatrix} 01011 \\ 11010 \end{bmatrix}$. Then $\hat{\theta}(\begin{bmatrix} 01011 \\ 11010 \end{bmatrix})$ contains the irreducible balanced block $\begin{bmatrix} 01001011 \\ 11010010 \end{bmatrix}$, and the $\hat{\theta}$ -image of this block contains a irreducible balanced block of length 27. One can show that arbitrary long irreducible balanced blocks will occur. In particular we do have a problem to compute δ_c for this θ .

3. BALANCING WITH WEIGHTS

To tackle the problem posed at the end of the previous section we generalize the notion of balanced blocks.

A *weight vector* $w^T = (w_0, w_1)$ consists of two positive reals w_0, w_1 . A block $\hat{A} = \begin{pmatrix} A \\ B \end{pmatrix}$ consisting of two words A and B is called a *w-balanced block* if

$$N_0(A)w_0 + N_1(A)w_1 = N_0(B)w_0 + N_1(B)w_1,$$

which we can shortly write as $N(A)^T w = N(B)^T w$.

Observe that this generalizes the notion of balancedness in the sense that \hat{A} balanced implies \hat{A} is w -balanced, irrespective of the choice of w . Throughout the paper the eigenvalues of M_θ will be denoted by λ° and λ , where λ is the eigenvalue with the largest modulus, also called PF eigenvalue.

Lemma 1. *Let θ be a primitive substitution with matrix M_θ . Let w be a Perron Frobenius right eigenvector of M_θ : $M_\theta w = \lambda w$. Then \hat{A} w -balanced $\Leftrightarrow \hat{\theta}(\hat{A})$ w -balanced.*

Proof: Recall that $M_\theta(i, j)$ equals the number of times the letter j occurs in $\theta(i)$. Thus

$$N(\theta(A))^T w = N(A)^T M_\theta w = \lambda N(A)^T w = \lambda N(B)^T w = N(B)^T M_\theta w = N(\theta(B))^T w. \quad \square$$

It will be useful for us to have an explicit expression for the Perron Frobenius right eigenvector. We denote by $|w|$ the length of a word w .

Lemma 2. *Let θ be a primitive substitution with matrix M_θ . Let λ°, λ with $|\lambda^\circ| < \lambda$ be the eigenvalues of M_θ . Then $w^T = (|\theta(0)| - \lambda^\circ, |\theta(1)| - \lambda^\circ)$ is a right eigenvector for λ .*

Proof: By Cayley Hamilton we obtain $(M_\theta - \lambda Id)(M_\theta - \lambda^\circ Id) \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. \square

We continue with the substitution $\theta : 0 \mapsto 010 \quad 1 \mapsto 11010$; here $M_\theta = \begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}$. A right eigenvector w is $w^T = (1, 2)^T$. So the ‘double’ sequence $\Delta^* = \begin{pmatrix} u^* \\ v^* \end{pmatrix}$ starts with the irreducible w -balanced block

$$\hat{C} = \left| \begin{array}{cc} 0 & 1 & 0 \\ 1 & 1 & \end{array} \right| \quad \text{directly followed by} \quad \hat{D} = \left| \begin{array}{cc} 1 & 1 \\ 0 & 1 & 0 \end{array} \right|.$$

By Lemma 1, we can decompose the $\hat{\theta}$ -image of these two blocks in irreducible w -balanced blocks:

$$\hat{\theta}(\hat{C}) = \left| \begin{array}{cc} 0 & 1 & 0 \\ 1 & 1 & 0 \end{array} \right| \left| \begin{array}{cc} 1 & 1 \\ 0 & 1 & 0 \end{array} \right| \left| \begin{array}{cc} 0 & 1 & 0 \\ 1 & 1 & 0 \end{array} \right| \left| \begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array} \right|, \quad \hat{\theta}(\hat{D}) = \left| \begin{array}{cc} 1 & 1 \\ 0 & 1 & 0 \end{array} \right| \left| \begin{array}{cc} 0 & 1 & 0 \\ 1 & 1 & 0 \end{array} \right| \left| \begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array} \right| \left| \begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array} \right|.$$

Note the occurrence of the ‘skewed’ w -balanced blocks $\hat{0}_{\natural} =: \left| \begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array} \right|$ and $\hat{1}_{\natural} =: \left| \begin{array}{cc} 1 & 1 \\ 0 & 1 \end{array} \right|$.

We see that Δ^* is generated by the substitution $\hat{\theta}$ given by:

$$\begin{aligned} \hat{\theta}(\hat{0}) &= \hat{0} \hat{1} \hat{0}, & \hat{\theta}(\hat{1}) &= \hat{1} \hat{1} \hat{0} \hat{1} \hat{0}, & \hat{\theta}(\hat{0}_{\natural}) &= \hat{0}_{\natural} \hat{1}_{\natural} \hat{0}_{\natural}, & \hat{\theta}(\hat{1}_{\natural}) &= \hat{1}_{\natural} \hat{1}_{\natural} \hat{0}_{\natural} \hat{1}_{\natural} \hat{0}_{\natural}, \\ \hat{\theta}(\hat{C}) &= \hat{C} \hat{D} \hat{C} \hat{0}_{\natural} \hat{1}_{\natural} \hat{0}_{\natural}, & \hat{\theta}(\hat{D}) &= \hat{D} \hat{C} \hat{D} \hat{0} \hat{1} \hat{0}. \end{aligned}$$

However,... we tacitly assumed that the blocks keep their ‘shape’! Actually, if $A = a_1 \dots a_m$, and $B = b_1 \dots b_n$, then a w -balanced block has a shape as

$$\left| \begin{array}{c} a_1 \dots a_m \\ b_1 \dots b_s b_{s+1} \dots b_{s+m} b_{s+m+1} \dots b_{n-1} b_n \end{array} \right|.$$

So a w -balanced block has two more parameters, a *left slope* s and a *right slope* $m - n + s$.

By definition the left slope of $\hat{\theta}(\hat{A})$ is equal to the left slope of \hat{A} .

Lemma 3. *Let θ be a primitive substitution of Salem type. Let $\hat{A} = \begin{pmatrix} A \\ B \end{pmatrix}$ be a w -balanced block. Then the right slope of $\hat{\theta}(\hat{A})$ is equal to the right slope of \hat{A} .*

Proof: Suppose that $\lambda^\circ = 1$, and let s be the left slope of \hat{A} . We have to show that $|\theta(A)| - |\theta(B)| + s = |A| - |B| + s$, which holds if and only if $|\theta(A)| - |A| = |\theta(B)| - |B|$. The left hand side equals

$$\begin{aligned} |\theta(A)| - |A| &= N_0(A)|\theta(0)| + N_1(A)|\theta(1)| - N_0(A) - N_1(A) \\ &= N_0(A)(|\theta(0)| - 1) + N_1(A)(|\theta(1)| - 1) = N(A)^T w, \end{aligned}$$

where w is a Perron Frobenius right eigenvector of M_θ (using Lemma 2). Since $\begin{pmatrix} A \\ B \end{pmatrix}$ is balanced $N(A)^T w = N(B)^T w$, and it follows that $|\theta(A)| - |A| = |\theta(B)| - |B|$.

In case $\lambda^\circ = -1$, one replaces θ by θ^2 , and proceeds as above. \square

4. COMPUTING THE COINCIDENCE DENSITY

The following general lemma will be useful.

Lemma 4. *Let (a_n) be a sequence of real numbers such that $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n a_i = a$ exists. Let (n_k) and (t_k) be two sequences with $t_k \geq 0$, $n_k \rightarrow \infty$ and $\limsup n_k/t_k < \infty$. Then*

$$\lim_{k \rightarrow \infty} \frac{1}{t_k} \sum_{i=n_k+1}^{n_k+t_k} a_i = a.$$

Proof: Let $\varepsilon > 0$. For all large n we will have $n(a - \varepsilon) \leq \sum_{i=1}^n a_i \leq n(a + \varepsilon)$. So

$$(n_k + t_k)(a - \varepsilon) - n_k(a + \varepsilon) \leq \sum_{i=1}^{n_k+t_k} a_i - \sum_{i=1}^{n_k} a_i \leq (n_k + t_k)(a + \varepsilon) - n_k(a - \varepsilon),$$

which is equivalent to

$$t_k a - \varepsilon(t_k + 2n_k) \leq \sum_{i=n_k+1}^{n_k+t_k} a_i \leq t_k a - \varepsilon(t_k + 2n_k).$$

Dividing through by t_k , and letting $k \rightarrow \infty$, the result follows by next letting $\varepsilon \downarrow 0$. \square

Our main result is

Theorem 1. *Let θ be a primitive substitution on $\{0,1\}^*$ of Salem type with $\theta(0) = 0\dots$, and $\theta(1) = 1\dots$. Let w be a Perron-Frobenius right eigenvector of M_θ . Suppose the set of w -balanced blocks is finite, and contains both $\hat{0}$ and $\hat{0}_{\frac{1}{2}}$. Then the coincidence density of θ does not exist.*

Proof: Suppose that δ_c does exist. Let j_0 be a position in the double sequence \hat{x} where $\hat{0}$ does occur. Let $A = x_1 \dots x_{j_0-1}$. It follows from Lemma 3 that for all $k \geq 1$ the block $\hat{\theta}^k(\hat{0})$ does occur in \hat{x} at position $n_k := |\theta^k(A)|$. Let $t_k := |\theta^k(0)|$. From the Perron-Frobenius theorem one deduces that $\limsup n_k/t_k < \infty$. We can therefore apply Lemma 4, and since we have t_k coincidences in the block $\hat{\theta}^k(\hat{0})$, it follows that $\delta_c = 1$. Similarly, let j_1 be a position where $\hat{0}_{\frac{1}{2}}$ occurs, and take n_k (with j_0 replaced by j_1) as before. Ignoring the lower left and the upper right letter of $\hat{\theta}^k(\hat{0}_{\frac{1}{2}})$, take $t_k = |\theta^k(0)| - 2$. It remains to determine the number of coincidences in the block $\hat{\theta}^k(\hat{0}_{\frac{1}{2}})$ (with the two letters removed). For any k this block contains a positive fraction of blocks $\hat{\theta}(\hat{0}_{\frac{1}{2}})$, so not to have a contradiction with $\delta_c = 1$, no *non*-coincidences are allowed for $\hat{\theta}(\hat{0}_{\frac{1}{2}})$. This requirement leads to $\theta(0) = 00\dots 0$, which contradicts primitivity of θ . Conclusion: δ_c does *not* exist. \square

Applying this result to the substitution $\theta : 0 \mapsto 010, 1 \mapsto 11010$ from the previous section, we find that the coincidence density does not exist here, nor will the frequencies of letters on the diagonal $\Delta^*(0,1)$ exist.

5. SELF-SIMILARITY CAN INCREASE REGULARITY

As in [13] one obtains a *self similar* 1D tiling from a substitution θ by associating to each letter i an interval of length w_i , where w is a right Perron-Frobenius eigenvector of the substitution matrix. This induces a 2D self-similar tiling from the direct product $\theta \times \theta$, by associating to each pair of letters (i,j) the $w_i \times w_j$ rectangle. Let us reconsider our example $\theta : 0 \mapsto 010, 1 \mapsto 11010$, choosing $w^T = (1,2)^T$. The diagonal of the 2D self-similar tiling associated to this example can be generated by a substitution δ on eight letters, corresponding to the ways in which the diagonal intersects a tile: a for the 1×1 tile, b and c for the 1×2 tile, etc., see Figure 3. The substitution δ is given by

$$\delta(a) = \delta(c) = \delta(e) = a h a, \delta(b) = \delta(f) = b g f c, \delta(d) = \delta(g) = d f g e, \delta(h) = h h a h a.$$

We can split the word $\delta(h)$ into the words $h h$ and $a h a$, and the letter h into two letters $h = \check{h}\check{h}$, obtaining a new substitution $\dot{\delta}$ on nine letters generating the diagonal, given by

$$\dot{\delta}(a) = \dot{\delta}(c) = \dot{\delta}(e) = \dot{\delta}(\check{h}) = a \check{h} \check{h} a, \dot{\delta}(b) = \dot{\delta}(f) = b g f c, \dot{\delta}(d) = \dot{\delta}(g) = d f g e, \dot{\delta}(\check{h}) = \check{h} \check{h} \check{h} \check{h}.$$

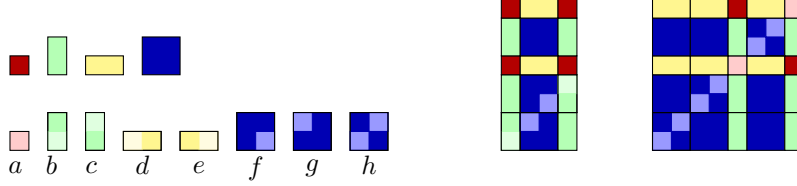


FIGURE 3. Four tiles, 8 diagonal tiles, and the first iterate of the b -tile and the h -tile.

Now note that δ is a substitution of constant length 4, and discrete. This implies that the frequency of letters in $\Delta^*(0, 1) = \Delta^*(b)$ exists (the letters a and h will have frequencies $2/3$ and $1/3$, the others frequency 0). This phenomenon is a special case of the following general result.

Proposition 1. *Let θ be a primitive substitution on $\{0, 1\}^*$ with $\theta(0) = 0\dots$, $\theta(1) = 1\dots$, for which M_θ has rational eigenvalues. Then the frequencies of letters on the diagonal of the 2D self-similar tiling generated by $\theta \times \theta$ do exist.*

Proof: Let λ be the PF eigenvalue of θ . Since sum and product of the eigenvalues are integers, it follows that λ is an integer, and thus there is an integer eigenvector, say $(p, q)^T$, with $p, q > 0$. Ignoring a scaling, we can assume that the 1D self-similar tiling generated by θ is built out of two tiles with length p , respectively q . The strategy is to cut up these two tiles in intervals of length 1, so that we may describe the diagonal of the direct product by the corresponding 1×1 squares (cf. the example above). Let a_k , $k = 1, \dots, p$ and b_ℓ , $\ell = 1, \dots, q$ be the letters of an alphabet \dot{A} of size $p + q$. Define the map $\gamma : \{0, 1\}^* \rightarrow \dot{A}^*$ by

$$\gamma(0) = a_1 a_2 \dots a_p, \quad \gamma(1) = b_1 b_2 \dots b_q.$$

Next, define the substitution $\dot{\theta} : \dot{A}^* \rightarrow \dot{A}^*$ by

$$\dot{\theta}(a_k) = \gamma(\theta(0))[(k-1)\lambda + 1, k\lambda], \quad k = 1, \dots, p, \quad \dot{\theta}(b_\ell) = \gamma(\theta(1))[(\ell-1)\lambda + 1, \ell\lambda], \quad \ell = 1, \dots, q.$$

Here we use the notation $w[m, n] := w_m \dots w_n$, and we remark that such a definition makes sense because of the eigenvalue equation $M_\theta \begin{pmatrix} p \\ q \end{pmatrix} = \lambda \begin{pmatrix} p \\ q \end{pmatrix}$. Now note that the diagonal of the direct product self-similar tiling is generated by $\dot{\theta} \times \dot{\theta}$, where $\dot{\theta}$ is a *constant length* substitution. As remarked in Section 1, this implies that the frequencies of letters exist. \square

6. ALL EIGENVALUES OUTSIDE THE UNIT CIRCLE

The case where the modulus of both eigenvalues of M_θ is larger than 1 is more difficult. It is conjectured in [11] that in this case the coincidence density will not exist.

First we discuss the case where the eigenvalues of M_θ are rational. As in the proof of Proposition 1 this implies that the smallest eigenvalue λ° is an integer. Let w be a right PF vector of M_θ . The set of w -balanced blocks will be either empty or infinite. This follows from an extension of Lemma 3: the left and the right slope of w -balanced blocks are multiplied by λ° when passing from \dot{A} to $\dot{\theta}(\dot{A})$. Thus the technique of w -balanced blocks is no longer useful. However, for the self-similar 2D tilings generated by θ Proposition 1 applies.

Secondly we consider irrational eigenvalues. With Lemma 2 one deduces from this irrationality that the only w -balanced blocks (with the PF weight vector w) are the balanced blocks. Using the condition $|\lambda^\circ| > 1$, one can show that the set of balanced blocks is either empty or infinite.

As an example of a substitution with irrational eigenvalues we consider

$$\theta : 0 \mapsto 00001, \quad 1 \mapsto 1110,$$

which has eigenvalues $\frac{7}{2} - \frac{1}{2}\sqrt{5}$ and $\frac{7}{2} + \frac{1}{2}\sqrt{5}$.

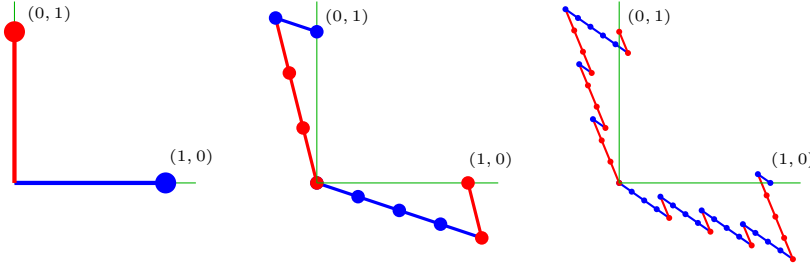


FIGURE 4. The first three approximants to the two curves $K(0)$ and $K(1)$ generated by the full representation of the substitution $0 \mapsto 00001$, $1 \mapsto 1110$.

Proposition 2. *The set of balanced blocks of the substitution θ given by $0 \mapsto 00001$, $1 \mapsto 1110$ is empty.*

Proof: We give a geometrical proof. We consider a geometric realization of θ given by its full representation, cf. [4]. This means that we consider the morphism $f : \{0,1\}^* \rightarrow \mathbb{R}^2$ given by $f(0) = (1,0)^T$, $f(1) = (0,1)^T$, and the linear map L_θ with matrix M_θ . Then we have the representation equation

$$f \circ \theta = L_\theta \circ f.$$

From [4] we have that curves

$$K_n(i) := L_\theta^{-n} K[\theta^n(i)]$$

converge exponentially fast to a curve $K(i)$ for $i = 0, 1$. Here for a word $w = w_1 \dots w_m$ we denote by $K[w]$ the polygonal curve passing successively through the points $(0,0), f(w_1), f(w_1 w_2), \dots, f(w_1 \dots w_m)$.

In Figure 4 we display $K_n(0)$ and $K_n(1)$ for $n = 0, 1, 2$. From the exponential convergence (speed $1/11$), and the fact that these curves (almost) live outside the positive quadrant, it is obvious that they will never intersect, and thus there is no balanced block. \square

The self similar 2D-tiling associated to the substitution $0 \mapsto 00001$, $1 \mapsto 1110$ is illustrated with Figure 5. The scalings are given by the eigenvector $(\frac{1}{2} + \frac{1}{2}\sqrt{5}, 1)$. We conjecture that there are infinitely many ‘types’ of squares and rectangles on the diagonal. At least it is clear that there are no ‘regeneration points’ as in the Pisot case, because of Proposition 2: the only $(\frac{1}{2} + \frac{1}{2}\sqrt{5}, 1)$ -balanced blocks are the balanced blocks, and these do not exist.

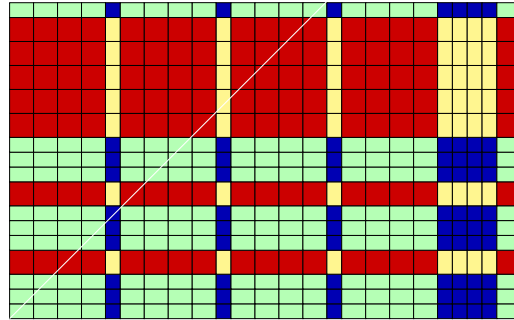


FIGURE 5. Part of the self-similar product tiling generated by the substitution $0 \mapsto 00001$, $1 \mapsto 1110$.

7. OVERLAP DISTRIBUTIONS

The coincidence density δ_c can be considered as a special case of the so called *overlap* between two sequences x and y defined by

$$\delta_c(x, y) := \lim_{n \rightarrow \infty} \frac{1}{n} \text{Card}\{1 \leq k \leq n : x_k = y_k\}.$$

Suppose x and y come from a shift invariant set X of sequences, equipped with a shift invariant measure μ . Then the distribution of $\delta_c(\cdot, \cdot)$ under the measure $\mu \times \mu$ is called the *overlap distribution*, introduced by Parisi ([10]). It was shown in [5] that if μ is weakly mixing, then the overlap distribution degenerates to a point mass concentrated on $\mu([0])^2 + \dots + \mu([r])^2$, where r is the number of letters in the alphabet, and $[i]$ is the cylinder $\{x \in X : x_0 = i\}$. (This follows immediately from the ergodic theorem applied with $\mu \times \mu$.)

Primitive substitutions θ generate a shift dynamical system with a unique invariant measure μ_θ . The overlap distribution of some of these is determined in [5] (see also [6]). For the substitution $\theta : 0 \mapsto 010 \quad 1 \mapsto 11010$ studied in Section 3, it is known that μ_θ is weak mixing ([9], [8]). So for this substitution $\delta_c(x, y)$ exists for μ_θ -almost all x and y , and for those x and y

$$\delta_c(x, y) = \mu_\theta([0])^2 + \mu_\theta([1])^2 = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}.$$

By Theorem 1 this means that the pair (u^*, v^*) is non-generic for the overlap distribution. In fact most substitutions with matrix $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ and two fixed points will have non-generic pairs (u^*, v^*) by an application of this theorem. One exception is the substitution given by $\theta(0) = 010, \theta(1) = 10101$, since it generates a periodic sequence. Another is the ‘Kakutani’ substitution (cf. [3]) given by

$$\theta(0) = 001, \quad \theta(1) = 11001.$$

Here one can show easily that $\delta_c(u^*, v^*) = \delta_c = \frac{1}{2}$, so in this case the pair (u^*, v^*) is generic.

8. FINAL REMARKS

We would like to point out that the phenomenon of non-existence of frequencies of symbols can already occur for substitutions of constant length. A simple example is given by

$$1 \mapsto 123, \quad 2 \mapsto 222, \quad 3 \mapsto 333.$$

Of the 12 substitutions with matrix $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ (and $\theta(0) = 0\dots, \theta(1) = 1\dots$) there are 9 for which the coincidence density does not exist. Then there is a periodic one, and the ‘Kakutani’ substitution mentioned in the previous section. There is one substitution left, given by $\theta(0) = 001, \theta(1) = 10110$. We do not know how to determine the coincidence density or frequencies of letters on the diagonal for this substitution. Nor do we know how to compute these for substitutions whose matrix has two eigenvalues outside the unit circle.

REFERENCES

- [1] Marcy Barge and Beverly Diamond. Coincidence for substitutions of Pisot type. *Bull. Soc. Math. France*, 130(4): 619–626, 2002.
- [2] Ethan M. Coven and Michael S. Keane. The structure of substitution minimal sets. *Trans. Amer. Math. Soc* 162: 89–102, 1971.
- [3] F. M. Dekking and M. Keane. Mixing properties of substitutions. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 42(1): 23–33, 1978.
- [4] F. M. Dekking. Recurrent sets. *Adv. Math.* 44: 78–104, 1982.
- [5] A. C. D. van Enter, A. Hof, and Jacek Mipolhkeisz. Overlap distributions for deterministic systems with many pure states. *J. Phys. A*, 25(18): L1133–L1137, 1992.
- [6] Aernout C. D. van Enter and Ellis de Groote. An ultrametric state space with a dense discrete overlap distribution: paperfolding sequences. *J. Stat. Phys.*, 142(2): 223–228, 2011.
- [7] Michael Hollander and Boris Solomyak. Two-symbol Pisot substitutions have pure discrete spectrum. *Ergodic Theory Dynam. Systems*, 23(2): 533–540, 2003.
- [8] B. Host. Valeurs propres des systèmes dynamiques définis par des substitutions de longueur variable. *Ergodic Theory Dynam. Systems*, 6(4): 529–540, 1986.

- [9] John C. Martin. Minimal flows arising from substitutions of non-constant length. *Math. Systems Theory*, 7: 72–82, 1973.
- [10] Parisi G. Order parameter for spin-glasses. *Phys. Rev. Lett.* 50: 1946–1948, 1983.
- [11] Pierre Michel. Coincidence values and spectra of substitutions. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 42: 205–227, 1978.
- [12] N.P. Frank, A primer of substitution tilings of the Euclidean plane, *Expo. Math.* 26: 295–326, 2008.
- [13] B. Solomyak, Dynamics of self-similar tilings, *Ergodic Theory Dynam. Systems* 17: 695–738, 1997.